

REMARKS

The present application was filed on March 19, 2004, with claims 1-25. Claims 1-25 remain pending, including independent claims 1, 14, 17, 18 and 25.

Claims 1, 5-9 and 14-17 are rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent Application Publication No. 2004/0162901 (hereinafter “Mangipudi”) in view of U.S. Patent Application Publication No. 2005/0198200 (hereinafter “Subramanian”).

Claims 2-4, 18-20 and 25 are rejected under 35 U.S.C. §103(a) as being unpatentable over Mangipudi and Subramanian in view of U.S. Patent No. 6,112,221 (hereinafter “Bender”) and U.S. Patent Application Publication No. 2003/0120705 (hereinafter “Chen”).

Claims 10-12 are rejected under 35 U.S.C. §103(a) as being unpatentable over Mangipudi and Subramanian in view of U.S. Patent No. 6,981,029 (hereinafter “Menditto”).

Claim 13 is rejected under 35 U.S.C. §103(a) as being unpatentable over Mangipudi, Subramanian and Menditto in view of U.S. Patent No. 6,772,211 (hereinafter “Lu”).

Claims 21-23 are rejected under 35 U.S.C. §103(a) as being unpatentable over Mangipudi, Subramanian, Bender, Chen and Menditto.

Claim 24 is rejected under 35 U.S.C. §103(a) as being unpatentable over Mangipudi, Subramanian, Bender, Chen, Menditto and Lu.

Applicants propose amending claim 1 without prejudice to clarify the recitation of scheduling submission of the request to the at least one server based on: (i) a quality-of-service (QoS) class assigned to a client from which the request originated; (ii) a response target associated with the QoS class; and (iii) an estimated response time associated with the at least one server, wherein scheduling submission of the request to the at least one server comprises determining when to submit the request to the at least one server. As amended, claim 1 recites determining when to submit the request to the at least one server based on: (i) a quality-of-service (QoS) class assigned to a client from which the request originated; (ii) a response target associated with the QoS class; and (iii) an estimated response time associated with the at least one server.

It is believed that this minor clarifying amendment merely clarifies the limitations already present in claim 1, and thus places the application in better form for consideration on appeal without

requiring further search or consideration. Entry of the amendment under 37 CFR 1.116(b)(2) is respectfully requested.

Support for the amendment may be found in the specification at, e.g., page 7, line 24, to page 8, line 1, which describes an illustrative embodiment in which “the scheduler typically needs to decide for each incoming request whether to admit the request immediately or to hold it back (withhold request). Moreover, the scheduler typically has to decide for requests that are held back when and in which order to admit them to the back-end server.” See also the specification at page 12, lines 8-21, with reference to FIG. 2.

In arguing that Mangipudi teaches scheduling submission of the request to the at least one server based on the factors recited in claim 1, the Examiner relies on Mangipudi at paragraph [0047], which describes “load sharing algorithms implemented in this illustrative embodiment to assign requests to back-end servers within a cluster. . . . All client requests sent to the illustrative embodiment are routed to the server selected as the most available and/or efficient server within each class according to a selected load balancing algorithm.” It is important to note that the load balancing algorithm in Mangipudi is being used solely to select which server client requests should be routed to, without determining when those requests should be submitted to the selected server.

Indeed, the Examiner concedes that Mangipudi does not teach determining when those requests should be submitted to the at least one server. Rather, the Examiner argues that “scheduling a task refers to the timing of performing a task.” Paragraph [0047] of Mangipudi does not describe an algorithm to be used for “scheduling,” but rather describes “load sharing algorithms . . . to assign requests to back-end servers within a cluster.”

The Examiner also describes paragraphs [0009] and [0011] of Mangipudi as disclosing “a well known technique of scheduling HTTP requests by placing them into queues [which] are serviced by the request controller based on configured policy such as length of queues, etc.” This placement into queues affects in what order the requests will be serviced by the request controller but it does not determine when the requests should be submitted to a server in the manner recited in claim 1. Moreover, it should be noted that the relied-upon portion of the background of Mangipudi

describes an entirely different product from that described in paragraph [0047] and which does not include any scheduling based on the three factors recited in claim 1.

Subramanian fails to remedy Mangipudi so as to teach or suggest determining when to submit the request to the at least one server based on: (i) a quality-of-service (QoS) class assigned to a client from which the request originated; (ii) a response target associated with the QoS class; and (iii) an estimated response time associated with the at least one server.

With regard to the §103 rejection of claims 18 and 25, Applicants note that the Examiner concedes that Mangipudi fails to disclose the limitation of claim 18 directed to withholding submission of a request when the request originated from a client assigned to a first QoS class to allow a request that originated from a client assigned to a second QoS class to meet a response target associated therewith. Indeed, as discussed above, Mangipudi only discloses determining which server a request should be submitted to, not when a request should be submitted to that server, and thus fails to teach or request any withholding submission of a request.

Rather, the Examiner argues that Bender discloses the limitation at issue. It is important to note that Bender deals exclusively with scheduling execution of jobs which either have already been submitted to the server (which Bender refers to as “on-line” scheduling) or which will be submitted at a definite future time (which Bender refers to as “off-line” scheduling). See Bender at column 3, lines 23-35. Bender does not teach or suggest any technique which involves withholding submission of requests to the server. Chen similarly does not involve withholding submission of requests to a server.

Moreover, as explained at column 5 of Bender:

At step 108, once the deadline for each uncompleted job is calculated, server system 10 schedules the jobs in accordance with an earliest deadline first ("EDF") methodology. With an EDF methodology, the first job that server system 10 schedules is the job which has the earliest deadline, as found in step 106, relative to all of the other jobs. It then chooses the job with the next earliest deadline, and schedules it second, and so on until all of the jobs have been scheduled.

At decision step 110, server system 10 inquires whether each and every one of the jobs have completion times which is earlier than each job's respective deadline, as found in step 106. If any job is not able to be completed prior to its deadline, then the estimated stretch

value is not feasible and is therefore adjusted at step 112. From step 112, the feasibility of the adjusted stretch value is re-checked by returning to step 106.

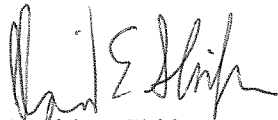
Thus, Bender schedules a job based on completion times and deadline times associated with *that particular job*. Bender does not withhold the request from submission to the at least one server when the request originated from a client assigned to a first QoS class to allow a request that originated from a client assigned to a second QoS class to meet a response target associated therewith. That is, Bender does not withhold a job based on a “response target” associated with *a particular QoS class*. Chen fails to remedy this fundamental deficiency of Mangipudi and Bender. Thus, Bender and Chen fail to remedy the admitted failure of Mangipudi to reach the limitations of claims 18 and 25.

In addition, Applicants, after considering the present Office Action in its entirety, respectfully assert the same deficiency arguments presented in their previous response dated May 5, 2008 (the disclosure of which is incorporated by reference herein) with respect to Veres, Menditto and Lu.

Applicants assert that the various dependent claims are not only patentable for the reasons given above but also because one or more of said claims recite separately patentable subject matter. For example, dependent claim 2 includes limitations similar to those discussed above with reference to claims 18 and 25 and hence is believed to be similarly patentable over the cited references.

In view of the above, Applicants believe that claims 1-25 are in condition for allowance, and again respectfully request withdrawal of the various remaining rejections.

Respectfully submitted,



David E. Shifren  
Attorney for Applicant(s)  
Reg. No. 59,329  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-2641

Date: May 17, 2010